

OM protein - protein search, using sw model

Run on: November 24, 2004, 13:52:15 : Search time 154 Seconds

(without alignments)  
1027.270 Million cell updates/sec

Title: US-09-513-151A-63

Perfect score: 441

Sequence: 1 MASVAAARAVFVSGSLRGLQ.....NKEPKGSGQNDLKCSV 441

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq\_23Sep04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	310	70.3	467	5	Abg96349 Human ova
2	274	62.1	465	5	Abg96348 Human ova
3	254	57.6	411	5	Abg97295 Novel hum
4	120	27.2	222	4	Aau23398 Novel hum
5	120	27.2	222	4	Abb10476 Human cdn
6	120	27.2	222	4	Abb10278 Human cdn
7	120	27.2	222	5	Abp67063 Human pol
8	120	27.2	222	5	Abp66865 Human pol
9	48	10.9	57	4	Aam15066 Peptide #
10	48	10.9	57	4	Abb34055 Peptide #
11	48	10.9	57	4	Aam27510 Peptide #
12	48	10.9	57	4	Abb28879 Peptide #
13	48	10.9	57	4	Aam67220 Human bon
14	48	10.9	57	4	Aam54837 Human bra
15	48	10.9	57	4	Abg48882 Human liv
16	16	3.6	221	2	Aay02534 Partial s
17	9	2.0	379	6	Ada34517 Acinetoba
18	9	2.0	387	7	Abc68188 Pseudomon
19	9	2.0	512	7	Aac23401 Pseudomon
20	8	1.8	68	8	Abc56542 Human gen
21	8	1.8	136	5	Abp31089 Human kin
22	8	1.8	139	5	Abp63871 Human ORF
23	8	1.8	200	7	Abc68993 Pseudomon
24	8	1.8	212	4	Aau65505 Propionib
25	8	1.8	212	6	Abm62024 Propionib

26	8	1.8	216	2	AAV37775	AAV37775 protein i
27	8	1.8	216	6	ABU27347	ABU27347 Protein e
28	8	1.8	219	4	AAU44657	AAU44657 Propionib
29	8	1.8	219	6	ABM41176	ABM41176 Propionib
30	8	1.8	248	7	ADH86448	ADH86448 Enterococ
31	8	1.8	257	6	ADB08980	ADB08980 Alloiooc
32	8	1.8	264	6	AAE30460	AAE30460 Haemophil
33	8	1.8	269	4	AAU65782	AAU65782 Propionib
34	8	1.8	269	6	ABM62301	ABM62301 Propionib
35	8	1.8	330	5	ABM92775	ABM92775 Herbicida
36	8	1.8	330	5	ABM93538	ABM93538 Herbicida
37	8	1.8	330	5	ABG70982	ABG70982 Arabidops
38	8	1.8	336	5	ABG70981	ABG70981 Arabidops
39	8	1.8	357	5	ABM91502	ABM91502 Herbicida
40	8	1.8	357	5	ABG70985	ABG70985 Arabidops
41	8	1.8	359	6	ABM72259	ABM72259 Staphyloc
42	8	1.8	419	6	ADA35828	ADA35828 Acinetoba
43	8	1.8	505	3	AAV69159	AAV69159 Peptide Q
44	8	1.8	568	6	ABM65635	ABM65635 Propionib
45	8	1.8	580	6	ABR53594	ABR53594 Protein s

#### ALIGNMENTS

RESULT 1  
ABG96349  
ID ABG96349 standard; protein; 467 AA.

XX ABG96349;

XX DT 11-DEC-2002 (first entry)

XX DE Human ovarian cancer marker M610.

XX KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;  
KW Central nervous system disorder; bacterial meningitis; viral meningitis;  
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;  
KW brain herniation; inflammation; encephalitis; testicular disorder;  
KW non-tuberculous granulomatous orchitis; connective tissue disorder;  
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;  
KW histological type; carcinogenic; ovarian cancer marker.

XX OS Homo sapiens.

XX WO200271928-A2.

XX PD 19-SEP-2002.

XX PD 14-MAR-2002; 2002WO-US007826.

XX PR 14-MAR-2001; 2001US-0276025P.

XX PR 14-MAR-2001; 2001US-0276026P.

XX PR 10-AUG-2001; 2001US-0311732P.

XX PR 19-SEP-2001; 2001US-0323580P.

XX PR 26-SEP-2001; 2001US-0324967P.

XX PR 26-SEP-2001; 2001US-0325102P.

XX PR 26-SEP-2001; 2001US-0325149P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;

XX PI Meyers RE, Morrissey MP, Glandt PJ, Sen A, Wieby PO, Mills GB;

XX PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;

XX WPI; 2002-723277/78.

XX N-PSDB; ABS76445.

XX Assessing whether a patient is afflicted with ovarian cancer, useful in assessing the stage or progression of the disease, comprises comparing the expression level of a cancer marker in a sample from a patient and from a non cancer patient.

PS Disclosure; Page 268-269; 481pp; English.

XX The present invention relates to a new method for assessing whether a patient is afflicted with ovarian cancer. The method involves comparing the expression level of a marker in a patient sample and the normal level of expression of the marker in a control non-ovarian cancer sample, where the marker is selected from 363 cancer markers described in the specification. The method of the invention is useful in diagnosing or characterizing cancer, in detecting the presence of cancer as early as possible, and the recurrence of ovarian cancer. The method may also be of particular use with patients having an enhanced risk of developing ovarian cancer (e.g. patients having a familial history of ovarian cancer). The cancer markers may be used in the management and treatment of e.g. brain and central nervous system disorders (e.g. bacterial and viral meningitis, Alzheimer's disease or Parkinson's disease), brain disorders (e.g. cerebral oedema, hydrocephalus or brain herniations), inflammations (e.g. bacterial or viral meningitis or encephalitis), testicular disorders (e.g. nontuberculous granulomatous orchitis), connective tissue disorders, or heart disorders (e.g. ischaemic heart disease or atherosclerosis). The compositions and methods may also be used in assessing the histological type of neoplasm associated with ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian carcinogenic potential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer. The present amino acid sequence represents one of the ovarian cancer markers described in the invention

XX

SQ Sequence 467 AA;

Query Match 70.3%; Score 310; DB 5; Length 467;  
Best Local Similarity 100.0%; Pred. No. 9.8e-304;  
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASVAAAPVPGSLRGQLRTPLVILGATGKSTLALQLGRLGGIVSADSMQVY 60

DB 1 MASVAAAPVPGSLRGQLRTPLVILGATGKSTLALQLGRLGGIVSADSMQVY 60

QY 61 EGLDIITNKVSAQQRICRHEMISFVDPVLTNTYVDFRNRATALIEDIFARDKIPIVWG 120

DB 61 EGLDIITNKVSAQQRICRHEMISFVDPVLTNTYVDFRNRATALIEDIFARDKIPIVWG 120

QY 121 GTNYYIESLLMKVLNTKQPMGTEKVIDRKLVELEKEDGLVHLKRLSQVDPPEMAKLUHPH 180

DB 121 GTNYYIESLLMKVLNTKQPMGTEKVIDRKLVELEKEDGLVHLKRLSQVDPPEMAKLUHPH 180

QY 181 DKRKVARSLQVFEETGISHSFLHROHTTEGGGPGGLGPKFSPNCIILWLHADOAVLDERL 240

DB 181 DKRKVARSLQVFEETGISHSFLHROHTTEGGGPGGLGPKFSPNCIILWLHADOAVLDERL 240

QY 241 DKRVDMLAAGLLEELRDFHRRYQKNVSENSQDYQHIGFOSIGFKFHEYLYTEGKCTL 300

DB 241 DKRVDMLAAGLLEELRDFHRRYQKNVSENSQDYQHIGFOSIGFKFHEYLYTEGKCTL 300

QY 301 ETSNQLKKG 310

DB 301 ETSNQLKKG 310

RESULT 2

ABG96348

XX ID ABG96348 standard; protein; 465 AA.

AC ABG96348;

XX

XX 11-DEC-2002 (first entry)

XX Human ovarian cancer marker OV73.

XX Human; ovarian cancer; marker: cancer; familial history; brain disorder;

XX Central nervous system disorder; bacterial meningitis; viral meningitis;

XX Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;

KW brain herniation; inflammation; encephalitis; testicular disorder;

KW nontuberculous granulomatous orchitis; connective tissue disorder;

KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;

XX histological type; carcinogenic; ovarian cancer marker.

OS Homo sapiens.

XX WO200271928-A2.

XX 19-SEP-2002.

XX 14-MAR-2002; 2002WO-US007826.

XX 14-MAR-2001; 2001US-0276025P.

XX 14-MAR-2001; 2001US-0276026P.

XX 10-AUG-2001; 2001US-0311732P.

XX 19-SEP-2001; 2001US-0323580P.

XX 26-SEP-2001; 2001US-0324967P.

XX 26-SEP-2001; 2001US-0325102P.

XX 26-SEP-2001; 2001US-0325149P.

XX (MILL-) MILLENNIUM PHARM INC.

PA Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;

PI Meyers RC, Morrisey MP, Olandt PJ, Sen A, Wieby PO, Mills GB;

PI East RE, Lu K, Schmandt RE, Zhao X, Glatt K;

XX WPI: 2002-723277/78.

DR N-PSDB; ABS76444.

XX

PT Assessing whether a patient is afflicted with ovarian cancer, useful in assessing the stage or progression of the disease, comprises comparing the expression level of a cancer marker in a sample from a patient and from a non cancer patient.

PT

XX Disclosure; Page 267-268; 481pp; English.

PS The present invention relates to a new method for assessing whether a patient is afflicted with ovarian cancer. The method involves comparing the expression level of a marker in a patient sample and the normal level of expression of the marker in a control non-ovarian cancer sample, where the marker is selected from 363 cancer markers described in the specification. The method of the invention is useful in diagnosing or characterizing cancer, in detecting the presence of cancer as early as possible, and the recurrence of ovarian cancer. The method may also be of particular use with patients having an enhanced risk of developing ovarian cancer (e.g. patients having a familial history of ovarian cancer). The cancer markers may be used in the management and treatment of e.g. brain and central nervous system disorders (e.g. bacterial and viral meningitis, Alzheimer's disease or Parkinson's disease), brain disorders (e.g. cerebral oedema, hydrocephalus or brain herniations), inflammations (e.g. bacterial or viral meningitis or encephalitis), testicular disorders (e.g. nontuberculous granulomatous orchitis), connective tissue disorders, or heart disorders (e.g. ischaemic heart disease or atherosclerosis). The compositions and methods may also be used in assessing the histological type of neoplasm associated with ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian carcinogenic potential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer. The present amino acid sequence represents one of the ovarian cancer markers described in the invention

XX

SQ Sequence 465 AA;

Query Match 62.1%; Score 274; DB 5; Length 465;  
Best Local Similarity 100.0%; Pred. No. 2.2e-267;  
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 STLALQRLGRLGGIVSADSMQVYEGGLIITNKVSAQQRICRHHMISFVDPVLTNTYV 96

DB 35 STLALQRLGRLGGIVSADSMQVYEGGLIITNKVSAQQRICRHHMISFVDPVLTNTYV 94

```
QY 97 DFNRRATALIEDIFARDKIPIVVGTTNYIESLLWKVLVNTKPOEMGTEKVIDRKVELEK 156
DB 95 DFNRRATALIEDIFARDKIPIVVGTTNYIESLLWKVLVNTKPOEMGTEKVIDRKVELEK 154
QY 157 EDGLVLHRLSVDPEMAAKLHPHDKRKVARS:QVFEETCISHSEFLHROHTBEGGGPLG 216
DB 155 EDGLVLHRLSVDPEMAAKLHPHDKRKVARS:QVFEETCISHSEFLHROHTBEGGGPLG 214
QY 217 GPLKFSNPCILWLHADQAVLDERLDKRVDDMLAAGLLEELDRFRRYKQNVNSQDYQ 276
DB 215 GPLKFSNPCILWLHADQAVLDERLDKRVDDMLAAGLLEELDRFRRYKQNVNSQDYQ 274
QY 277 HGIFQSIGKFEFHEYLITECKTLETNSQLLXKG 310
DB 275 HGIFQSIGKFEFHEYLITECKTLETNSQLLXKG 308

RESULT 3
ABB97295
XX AC ABB97295 standard; protein; 411 AA.
XX DT 27-JUN-2002 (first entry)
XX DE Novel human protein SEQ ID NO: 563.
XX KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
XX KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX KW neuroprotective; antiparkinsonian; protein therapy; EST;
XX KW expressed sequence tag.
XX OS Homo sapiens.
XX FN WO200222660-A2.
XX PD 21-MAR-2002.
XX PF 10-SEP-2001; 2001WO-US026015.
XX PR 11-SEP-2000; 2000US-00659671.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX DR WPI; 2002-292408/33.
XX DR N-PSDB; ABN32481.
XX PT An isolated polynucleotide for treating diseases associated with its
XX PT encoded polypeptide such as cancer and multiple sclerosis.
XX PS Example 2; SEQ ID NO 563; 509pp; English.
XX CC The present invention provides the protein and coding sequences of 444
XX CC novel human proteins. These were isolated from expressed sequences tags
XX CC (ESTs). They can be used to stimulate cell growth, to regulate
XX CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX CC e.g. in burn treatment, to regulate the immune system e.g. to treat
XX CC multiple sclerosis, to regulate activin or inhibit e.g. to treat
XX CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX CC rheumatoid arthritis, and to treat nervous system disorders e.g.
XX CC Parkinson's disease. The present sequence is a protein of the invention
XX SQ Sequence 411 AA;
```

```
Query Match 57.6%; Score 254; DB 5; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.1e-247;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 MOVYEGLDIITNKVSAQEQRICRHHMISFVDPPLVNTYTVDFRNRATALIEDIFARDKIP 116
DB 1 MOVYEGLDIITNKVSAQEQRICRHHMISFVDPPLVNTYTVDFRNRATALIEDIFARDKIP 60
QY 117 IVVGTTNYIESLLWKVLVNTKPOEMGTEKVIDRKVELEKEDGLVHRLSVDPEMAAK 176
DB 61 IVVGTTNYIESLLWKVLVNTKPOEMGTEKVIDRKVELEKEDGLVHRLSVDPEMAAK 120
QY 177 LHPHDKRKVARSLOVFEETGISHSEFLHROHTBEGGGPLGKFSNPCILWLHADQAVL 236
DB 121 LHPHDKRKVARSLOVFEETGISHSEFLHROHTBEGGGPLGKFSNPCILWLHADQAVL 180
QY 237 DERLDKRVDDMLAAGLLEELDRFRRYKQNVNSQDYQHGIQSIGKFEFHEYLITEG 296
DB 181 DERLDKRVDDMLAAGLLEELDRFRRYKQNVNSQDYQHGIQSIGKFEFHEYLITEG 240
QY 297 KCTLETNSQLLXKG 310
DB 241 KCTLETNSQLLXKG 254

RESULT 4
AAU23398
ID AAU23398 standard; protein; 222 AA.
XX AC AAU23398;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human enzyme polypeptide #484.
XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
XX KW ligase; hyperproliferative disorder; immunodeficiency disorder;
XX KW autoimmune disorder; neurological disorder; metabolic disorder;
XX KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
XX KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
XX KW nephrotropic; anticoagulant.
XX OS Homo sapiens.
XX FN WO200155301-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001239.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217498P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225211P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
```

PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0225779P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250191P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465566/50.  
DR N-PSDB; AAS41268.  
XX  
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,  
PT treating neural, immune system, muscular, reproductive, pulmonary,  
PT cardiovascular, renal, proliferative disorders and cancerous diseases.  
XX  
PS Claim 11; SEQ ID NO 1394; 1180pp; English.  
XX  
CC The present invention relates to the isolation of novel human enzyme  
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
CC isomerases or ligases. The sequences of the invention are useful in the  
CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
CC disorders including hyperproliferative disorders (e.g. cancer),  
CC immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g.  
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic  
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),  
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders  
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and  
CC infectious disorders (e.g. influenza). The polynucleotides of the  
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent  
CC the novel human enzyme polypeptides of the invention. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 222 AA;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MASVAARAVPVGSLRGQLRTPLVILGATGTGKSTLALQGLGRLGGEIVSADSMQVY 60
Db	4 MASVAARAVPVGSLRGQLRTPLVILGATGTGKSTLALQGLGRLGGEIVSADSMQVY 63
QY	61 EGLDIITNKVSAQQRICRHHMISFVDPLVNTVTVDFNRATALLIEDIPARDKIPVVG 120
Db	64 EGLDIITNKVSAQQRICRHHMISFVDPLVNTVTVDFNRATALLIEDIPARDKIPVVG 123
RESULT 5	
ABBI0476	
ID	ABBI0476 standard; protein; 222 AA.
XX	AC ABB10476;
XX	XX
DT	10-JAN-2002 (first entry)
XX	XX
DE	Human cDNA SEQ ID NO: 784.
XX	XX
KW	Human; gene therapy; neural disorder; immune system disorder;
KW	muscular disorder; reproductive disorder; gastrointestinal disorder;
KW	pulmonary disorder; cardiovascular disorder; renal disorder;
KW	proliferative disorder; inflammation.
XX	XX
OS	Homo sapiens.
XX	XX
FN	W0200154474-A2.
XX	XX
PD	02-AUG-2001.
XX	XX
PF	17-JAN-2001; 2001WO-US001349.
XX	XX
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216680P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218290P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225266P.
PR	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225447P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225759P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226686P.
PR	22-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.

PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0232401P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	14-SEP-2000; 2000US-0233065P.
PR	21-SEP-2000; 2000US-0234223P.
PR	21-SEP-2000; 2000US-0234274P.
PR	25-SEP-2000; 2000US-0234997P.
PR	25-SEP-2000; 2000US-0234998P.
PR	26-SEP-2000; 2000US-0235484P.
PR	27-SEP-2000; 2000US-0235834P.
PR	27-SEP-2000; 2000US-0235836P.
PR	29-SEP-2000; 2000US-0236327P.
PR	29-SEP-2000; 2000US-0236367P.
PR	29-SEP-2000; 2000US-0236368P.
PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	02-OCT-2000; 2000US-0237040P.
PR	13-OCT-2000; 2000US-0239935P.
PR	13-OCT-2000; 2000US-0239937P.
PR	20-OCT-2000; 2000US-0240960P.
PR	20-OCT-2000; 2000US-0241221P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT-2000; 2000US-0241808P.
PR	20-OCT-2000; 2000US-0241809P.
PR	20-OCT-2000; 2000US-0241826P.
PR	01-NOV-2000; 2000US-0244617P.
PR	08-NOV-2000; 2000US-0246474P.
PR	08-NOV-2000; 2000US-0246475P.
PR	08-NOV-2000; 2000US-0246476P.
PR	08-NOV-2000; 2000US-0246477P.
PR	08-NOV-2000; 2000US-0246478P.
PR	08-NOV-2000; 2000US-0246523P.
PR	08-NOV-2000; 2000US-0246524P.
PR	08-NOV-2000; 2000US-0246525P.
PR	08-NOV-2000; 2000US-0246526P.
PR	08-NOV-2000; 2000US-0246528P.
PR	08-NOV-2000; 2000US-0246529P.
PR	08-NOV-2000; 2000US-0246532P.
PR	08-NOV-2000; 2000US-0246609P.
PR	08-NOV-2000; 2000US-0246610P.
PR	08-NOV-2000; 2000US-0246611P.
PR	08-NOV-2000; 2000US-0246613P.
PR	17-NOV-2000; 2000US-0249207P.
PR	17-NOV-2000; 2000US-0249208P.
PR	17-NOV-2000; 2000US-0249209P.
PR	17-NOV-2000; 2000US-0249210P.
PR	17-NOV-2000; 2000US-0249211P.
PR	17-NOV-2000; 2000US-0249212P.
PR	17-NOV-2000; 2000US-0249213P.
PR	17-NOV-2000; 2000US-0249214P.

```

PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249222P.
PR 17-NOV-2000; 2000US-0249223P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249225P.
PR 17-NOV-2000; 2000US-0249226P.
PR 17-NOV-2000; 2000US-0249227P.
PR 17-NOV-2000; 2000US-0249228P.
PR 17-NOV-2000; 2000US-0249229P.
PR 17-NOV-2000; 2000US-0249230P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0255719P.
PR 06-DEC-2000; 2000US-0255719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0255678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI: 2001-476161/51.
DR N-PSDB; ABA06698.
XX
PT Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical
PT condition.
XX
PS Claim 11; SEQ ID NO 784; 859pp + Sequence Listing; English.
XX
CC The present invention provides human cDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a protein of the invention
XX
SQ Sequence 222 AA;
XX
Query Match 27.2%; Score 120; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.6e-112;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASVAAARAVPGSLRGLQRTLPVVLGATGKSTLALQLCRLGGHVSADSMQVY 60
DQ |||||
DQ 4 MASVAAARAVPGSLRGLQRTLPVVLGATGKSTLALQLCRLGGHVSADSMQVY 63
QY 61 EGLDIITNKVSAQQRICRHHMISFVDPPLVNTYTVDFRNRTALIEDIFARDKIPVVG 120
DQ |||||
DQ 64 EGLDIITNKVSAQQRICRHHMISFVDPPLVNTYTVDFRNRTALIEDIFARDKIPVVG 123
RESULT 6
ABBI0278
ID ABBI0278 standard; protein; 222 AA.
XX
AC ABBI0278;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human cDNA SEQ ID NO: 586.
XX
KW Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation.

```

```

XX Homo sapiens.
XX OS
XX XX WO200154474-A2.
XX PD
XX XX 02-AUG-2001.
XX PF
XX PF 17-JAN-2001; 2001WO-US001349.
XX PR
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225757P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 14-AUG-2000; 2000US-0225759P.
XX PR 18-AUG-2000; 2000US-0226779P.
XX PR 22-AUG-2000; 2000US-0226681P.
XX PR 22-AUG-2000; 2000US-0226686P.
XX PR 22-AUG-2000; 2000US-0227182P.
XX PR 23-AUG-2000; 2000US-0227009P.
XX PR 30-AUG-2000; 2000US-022824P.
XX PR 01-SEP-2000; 2000US-0228287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
XX PR 01-SEP-2000; 2000US-0229345P.
XX PR 05-SEP-2000; 2000US-0229509P.
XX PR 05-SEP-2000; 2000US-0229513P.
XX PR 06-SEP-2000; 2000US-0230437P.
XX PR 06-SEP-2000; 2000US-0230438P.
XX PR 08-SEP-2000; 2000US-0231242P.
XX PR 08-SEP-2000; 2000US-0231243P.
XX PR 08-SEP-2000; 2000US-0231244P.
XX PR 08-SEP-2000; 2000US-0231413P.
XX PR 08-SEP-2000; 2000US-0231414P.
XX PR 08-SEP-2000; 2000US-0232080P.
XX PR 08-SEP-2000; 2000US-0232081P.
XX PR 12-SEP-2000; 2000US-0231968P.
XX PR 14-SEP-2000; 2000US-0232397P.
XX PR 14-SEP-2000; 2000US-0232398P.
XX PR 14-SEP-2000; 2000US-0232399P.
XX PR 14-SEP-2000; 2000US-0232400P.
XX PR 14-SEP-2000; 2000US-0232401P.
XX PR 14-SEP-2000; 2000US-0233063P.
XX PR 14-SEP-2000; 2000US-0233064P.
XX PR 14-SEP-2000; 2000US-0233065P.
XX PR 21-SEP-2000; 2000US-0234223P.
XX PR 21-SEP-2000; 2000US-0234274P.
XX PR 25-SEP-2000; 2000US-0234997P.
XX PR 25-SEP-2000; 2000US-0234998P.

```



PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0237040P.  
 PR 20-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241803P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0245299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.  
 (RUBE/) RUBEN S M.  
 (BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2002-681727/73.

N-PSDB; ABV84035.

Novel polypeptide useful for diagnosis, prognosis, prevention, and treatment of immune, hyperproliferative, renal, respiratory, cardiovascular, reproductive, endocrine, gastrointestinal and neurological disorders.

Claim 11; SEQ ID NO 784; 369pp + Sequence Listing; English.

The invention relates to novel genes (ABV83682-ABV84101) and proteins (ABP66710-ABP67129) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing (e) neurological disorders such as e.g. cerebral ischaemia and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 222 AA;

Query Match 27.2%; Score 120; DB 5; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-112; Indels 0; Gaps 0;  
 Matches 120; Conservative 0; Mismatches 0;  
 QY 1 MASVAAARAVPVGSLRGRLQRTLPVVLGATGTGKSTLALQLGRLGGEIVSADSMQVY 60  
 Db 4 MASVAAARAVPVGSLRGRLQRTLPVVLGATGTGKSTLALQLGRLGGEIVSADSMQVY 63  
 QY 61 EGLDIITNKVSAQEQRICRHHMISFVDPLVTNTYVDFRNRTALIEDIFARDKIPVVG 120  
 Db 64 EGLDIITNKVSAQEQRICRHHMISFVDPLVTNTYVDFRNRTALIEDIFARDKIPVVG 123

# RESULT 8

ABP66865  
 ID ABP66865 standard; protein; 222 AA.

XX ABP66865;

XX 09-DEC-2002 (first entry)

XX Human polypeptide SEQ ID NO 586.

XX Human; neutropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
 KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;  
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX US2002090672-A1.

XX 11-JUL-2002.

XX 17-JAN-2001; 2001US-00764853.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216800P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 22-AUG-2000; 2000US-0225759P.

XX 30-AUG-2000; 2000US-0226868P.

XX 01-SEP-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 08-SEP-2000; 2000US-0229513P.

XX 21-SEP-2000; 2000US-0234223P.

XX 21-SEP-2000; 2000US-0234274P.

XX 25-SEP-2000; 2000US-0234997P.

XX 27-SEP-2000; 2000US-0235834P.

XX 29-SEP-2000; 2000US-0236327P.

XX 29-SEP-2000; 2000US-0236367P.



PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
DR WPI; 2002-681727/73.  
DR N-PSDB; ABV83837.  
XX  
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and  
PT treatment of immune, hyperproliferative, renal, respiratory, and  
PT cardiovascular, reproductive, endocrine, gastrointestinal and  
PT neurological disorders.  
XX  
PS Claim 11; SEQ ID NO 586; 369pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (ABV83682-ABV84101) and proteins  
CC (ABP6710-ABP67129) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 222 AA;  
Query Match 27.2%; Score 120; DB 5; Length 222;  
Best Local Similarity 100.0%; Pred. No. 3.6e-112; Mismatches 0; Gaps 0;  
Matches 120; Conservative 0; Indels 0; Indels 0; Gaps 0;  
QY 1 MASVAAARAVPVGSLGRLQRTLPVLVILGATGKSTLALQLQRLGGEIVSADSMQVY 60  
DB 4 MASVAAARAVPVGSLGRLQRTLPVLVILGATGKSTLALQLQRLGGEIVSADSMQVY 63  
QY 61 EGLDITITNKVSAQQRICRHHMISFVDPPLVNTYVDFRNRTALIEDIPARDKIPVWG 120  
DB 64 EGLDITITNKVSAQQRICRHHMISFVDPPLVNTYVDFRNRTALIEDIPARDKIPVWG 123  
RESULT 9  
ID AAM15066  
XX AAM15066 standard; protein; 57 AA.  
AC AAM15066;  
XX AAM15066;  
XX

DT 12-OCT-2001 (first entry)  
XX Peptide #1500 encoded by probe for measuring cervical gene expression.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
XX cervical cancer.  
KW Homo sapiens.  
XX  
OS WO200157278-A2.  
PN 09-AUG-2001.  
PD 30-JAN-2001; 2001WO-US000670.  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207458P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
DR Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human cervical epithelial cells.  
XX  
PS Claim 27; SEQ ID NO 19892; 487pp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENP; see AAI10068-AA128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 57 AA;  
Query Match 10.9%; Score 48; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 5.5e-40; Mismatches 0; Gaps 0;  
Matches 48; Conservative 0; Indels 0; Indels 0; Gaps 0;  
QY 59 VVEGLDITITNKVSAQQRICRHHMISFVDPPLVNTYVDFRNRTALI 106  
DB 1 VVEGLDITITNKVSAQQRICRHHMISFVDPPLVNTYVDFRNRTALI 48  
RESULT 10  
ID ABB34055  
XX ABB34055 standard; peptide; 57 AA.  
AC ABB34055;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #1561 encoded by human foetal liver single exon probe.  
XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX Homo sapiens.  
XX WO200157277-A2.  
XX 09-AUG-2001.  
XX

```

XX PF 30-JAN-2001; 2001WO-US000669.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human fetal liver.
XX PS Claim 27; SEQ ID NO 26690; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human fetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human fetal liver. The
XX CC present sequence is a peptide encoded by a single exon nucleic acid probe
XX CC of the invention. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 57 AA;

Query Match 10.9%; Score 48; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 5.5e-40;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VYEGLDITNKVSAQEQRICRHHMISFVDPLVTNTYVDFRNATALI 106
Db 1 VYEGLDITNKVSAQEQRICRHHMISFVDPLVTNTYVDFRNATALI 48

RESULT 11
AM27510
ID AM27510 standard; protein; 57 AA.
XX AC AM27510;
XX DT 17-OCT-2001 (first entry)
XX DE Peptide #1547 encoded by probe for measuring placental gene expression.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX DR New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 27; SEQ ID NO 11847; 327pp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and Bt 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each

```

---

```

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-48897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 27; SEQ ID NO 27779; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AA131315-AA157546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders
XX SQ Sequence 57 AA;

Query Match 10.9%; Score 48; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 5.5e-40;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VYEGLDITNKVSAQEQRICRHHMISFVDPLVTNTYVDFRNATALI 106
Db 1 VYEGLDITNKVSAQEQRICRHHMISFVDPLVTNTYVDFRNATALI 48

RESULT 12
ABB28879
ID ABB28879 standard; peptide; 57 AA.
XX AC ABB28879;
XX DT 01-FEB-2002 (first entry)
XX DE Peptide #1530 encoded by breast cell single exon nucleic acid probe.
XX KW Human; microarray; single exon probe; gene expression; breast; disease;
XX KW cancer.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000662.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX DR New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX PS Claim 27; SEQ ID NO 11847; 327pp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and Bt 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each

```

CC probe of the microarray. The probes are useful for verifying the  
CC expression of regions of genomic DNA predicted to encode proteins. They  
CC are useful for gene discovery, and for determining predisposition and/or  
CC prognosing breast disease. Gene expression analysis is useful for  
CC assessing the toxicity of chemical agents on cells. The microarray of  
CC this invention presents a far greater diversity of probes for measuring  
CC gene expression, with far less bias than expressed sequence tag  
CC microarrays. The method is suitable for rapid production of functional  
CC information from genomic sequence. The present sequence is a peptide  
CC encoded by a single exon nucleic acid probe of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 57 AA;

Query Match 10.9%; Score 48; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 5.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VYEGLDIITNKVSAQEQICRHHMISFVDPPLVTNTYVDFRNATALI 106  
DB 1 VYEGLDIITNKVSAQEQICRHHMISFVDPPLVTNTYVDFRNATALI 48

RESULT 13  
AAM67220  
ID AAM67220 standard; protein; 57 AA.  
XX  
AC AAM67220;

XX 06-NOV-2001 (first entry)  
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 27526.  
DE Human; bone marrow expressed exon; gene expression analysis; probe;  
XX microarray; cancer; leukaemia; lymphoma; myeloma.  
KW Homo sapiens.  
XX  
XX WO200157276-A2.  
XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488900/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human bone marrow.  
XX Example 4; SEQ ID NO 27526; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukaemia and myeloma. The present sequence is a  
XX protein encoded by one of the probes of the invention  
XX  
XX Sequence 57 AA;

Query Match 10.9%; Score 48; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 5.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VYEGLDIITNKVSAQEQICRHHMISFVDPPLVTNTYVDFRNATALI 106  
DB 1 VYEGLDIITNKVSAQEQICRHHMISFVDPPLVTNTYVDFRNATALI 48

RESULT 14  
AAM54837  
ID AAM54837 standard; protein; 57 AA.  
XX  
AC AAM54837;

XX 05-NOV-2001 (first entry)  
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 26942.  
DE Human; brain expressed exon; gene expression analysis; probe; microarray;  
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
KW Homo sapiens.  
XX  
XX WO200157275-A2.  
XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.  
XX 04-FEB-2000; 2000US-0180312P.  
XX 26-MAY-2000; 2000US-0207456P.  
XX 30-JUN-2000; 2000US-00608408.  
XX 03-AUG-2000; 2000US-00632366.  
XX 21-SEP-2000; 2000US-0234687P.  
XX 27-SEP-2000; 2000US-0236359P.  
XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483446/52.  
XX Single exon nucleic acid probes for analyzing gene expression in human  
XX brains.  
XX Example 4; SEQ ID NO 26942; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is a protein encoded by one of  
XX the probes of the invention  
XX  
XX Sequence 57 AA;

Query Match 10.9%; Score 48; DB 4; Length 57;  
Best Local Similarity 100.0%; Pred. No. 5.5e-40;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VYEGLDIITNKVSAQEQICRHHMISFVDPPLVTNTYVDFRNATALI 106  
DB 1 VYEGLDIITNKVSAQEQICRHHMISFVDPPLVTNTYVDFRNATALI 48

RESULT 15  
ABG48882  
ID ABG48882 standard; peptide; 57 AA.  
XX



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 24, 2004, 13:52:15 ; Search time 39 Seconds  
(without alignments)  
749.904 Million cell updates/sec

Title: US-09-513-151A-63

Perfect score: 441  
Sequence: 1 MASVAARAVPVGSLRGLQ.....NKEPKGKSGQNDLKC5V 441

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap.\*
- 5: /cgn2\_6/ptodata/1/iaa/6C COMB.pap.\*
- 6: /cgn2\_6/ptodata/1/iaa/6D COMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	3.6	252	4	US-09-270-767-44570
2	9	2.0	379	4	US-09-328-352-5804
3	9	2.0	386	4	US-09-248-796A-17815
4	9	2.0	387	4	US-09-252-991A-16934
5	9	2.0	512	4	US-09-716-865-14
6	8	1.8	200	4	US-09-252-991A-18739
7	8	1.8	248	4	US-09-134-000C-4333
8	8	1.8	419	4	US-09-328-352-7115
9	8	1.8	505	3	US-09-509-902A-11
10	8	1.8	580	4	US-09-538-092-40
11	8	1.8	661	4	US-09-107-532A-3677
12	8	1.8	666	3	US-09-134-001C-5465
13	8	1.8	702	4	US-09-328-352-4189
14	7	1.6	26	1	US-07-942-245-287
15	7	1.6	26	1	US-07-942-245-325
16	7	1.6	246	1	US-07-942-245-331
17	7	1.6	64	4	US-09-583-110-4731
18	7	1.6	74	4	US-09-489-039A-12687
19	7	1.6	90	4	US-09-248-796A-25818
20	7	1.6	104	2	US-09-609-324A-8
21	7	1.6	104	2	US-09-920-440B-8
22	7	1.6	104	3	US-09-173-432-8
23	7	1.6	104	3	US-09-173-133-8
24	7	1.6	104	3	US-09-165-533-8
25	7	1.6	104	4	US-09-580-236A-8
26	7	1.6	110	1	US-08-466-886-29
27	7	1.6	110	3	US-08-469-617-29

28 7 1.6 110 4 US-08-469-630-29 Sequence 29, Appl  
29 7 1.6 124 4 US-09-134-000C-4138 Sequence 4138, Ap  
30 7 1.6 129 6 5428135-6 Patent No. 5428135  
31 7 1.6 157 4 US-09-252-991A-24100 Sequence 24100, A  
32 7 1.6 158 4 US-09-489-039A-7655 Sequence 7655, Ap  
33 7 1.6 175 4 US-09-216-333B-6 Sequence 6, Appl  
34 7 1.6 184 1 US-09-609-324A-2 Sequence 2, Appl  
35 7 1.6 184 2 US-08-920-440B-2 Sequence 2, Appl  
36 7 1.6 184 3 US-09-173-492-2 Sequence 2, Appl  
37 7 1.6 184 3 US-09-173-133-2 Sequence 2, Appl  
38 7 1.6 184 3 US-09-165-533-2 Sequence 2, Appl  
39 7 1.6 184 4 US-09-580-236A-2 Sequence 2, Appl  
40 7 1.6 184 5 PCT-US95-12779-2 Sequence 2, Appl  
41 7 1.6 184 5 PCT-US95-15781-2 Sequence 2, Appl  
42 7 1.6 186 1 US-08-117-083-22 Sequence 22, Appl  
43 7 1.6 192 4 US-09-319-588C-8 Sequence 8, Appl  
44 7 1.6 209 1 US-08-455-001-2 Sequence 2, Appl  
45 7 1.6 209 3 US-08-308-814-2 Sequence 2, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-270-767-44570  
; Sequence 44570, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 44570  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-44570

Query Match 3.6%; Score 16; DB 4; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 116 PIVVGTYNYIESLLW 131  
Db 99 PIVVGTYNYIESLLW 114

##### RESULT 2

US-09-328-352-5804  
; Sequence 5804, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5804  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: *Acinetobacter baumannii*  
US-09-328-352-5804

Query Match 2.0%; Score 9; DB 4; Length 379;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 30 GATGTGKST 38

```
Db      138 GATGTGKST 146
|||||
RESULT 3
US-09-248-796A-17815
; Sequence 17815, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17815
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17815
Query Match      2.0%; Score 9; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      114 KPIVVVGST 122
|||||
Db      32 KPIVVVGST 40
|||||
RESULT 4
US-09-252-991A-16934
; Sequence 16934, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16934
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16934
Query Match      2.0%; Score 9; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 GATGTGKST 38
|||||
Db      135 GATGTGKST 143
|||||
RESULT 5
US-09-716-865-14
; Sequence 14, Application US/09716865
; Patent No. 6548292
; GENERAL INFORMATION:
; APPLICANT: Bramucci, Micheal G
; APPLICANT: Nagarajan, Vasantha
; APPLICANT: Chen, Mario W.
; TITLE OF INVENTION: Bacterial Plasmid Having Genes Encoding Enzymes for the
; TITLE OF INVENTION: Degradation of Aromatic Compounds
; FILE REFERENCE: BC1016 US NA
; CURRENT APPLICATION NUMBER: US/09/716,865
; CURRENT FILING DATE: 2000-11-26
; PRIOR APPLICATION NUMBER: 60/167,062
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Pseudomonas CT14
US-09-716-865-14
Query Match      2.0%; Score 9; DB 4; Length 512;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 GATGTGKST 38
|||||
Db      131 GATGTGKST 139
|||||
RESULT 6
US-09-252-991A-18739
; Sequence 18739, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18739
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18739
Query Match      1.8%; Score 8; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ASVAAARA 9
|||||
Db      170 ASVAAARA 177
|||||
RESULT 7
US-09-134-000C-4333
; Sequence 4333, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4333
; LENGTH: 248
```

```
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4333

Query Match      1.8%; Score 8; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      29 LGATGTGK 36
      |||||
Db      21 LGATGTGK 28

RESULT 8
US-09-328-352-7115
; Sequence 7115, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7115
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7115

Query Match      1.8%; Score 8; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 18;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      24 PLVVILGA 31
      |||||
Db      326 PLVVILGA 333

RESULT 9
US-09-509-902A-11
; Sequence 11, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-902A-11

Query Match      1.8%; Score 8; DB 3; Length 505;
Best Local Similarity 100.0%; Pred. No. 22;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 SVAAARAV 10
      |||||
Db      496 SVAAARAV 503

RESULT 10
US-09-538-092-40
; Sequence 40, Application US/09538092
```

```
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormat Version 0.9
; SEQ ID NO 40
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YER170C
US-09-538-092-40

Query Match      1.8%; Score 8; DB 4; Length 580;
Best Local Similarity 100.0%; Pred. No. 24;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      153 ELEKEDGL 160
      |||||
Db      121 ELEKEDGL 128

RESULT 11
US-09-107-532A-3677
; Sequence 3677, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3677:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661 amino acids
; TYPE: amino acid
```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...661
; SEQUENCE DESCRIPTION: SEQ ID NO: 3677:
US-09-107-532A-3677

Query Match          1.8%; Score 8; DB 4; Length 661;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LGATGTGK 36
DB 40 LGATGTGK 47

RESULT 12
US-09-134-001C-5465
; Sequence 5465, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5465
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5465

Query Match          1.8%; Score 8; DB 3; Length 666;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LGATGTGK 36
DB 45 LGATGTGK 52

RESULT 13
US-09-328-352-4189
; Sequence 4189, Application US/09328352
; Patent No. 6562558
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4189
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4189

Query Match          1.8%; Score 8; DB 4; Length 702;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 42 QLGQRLGG 49
DB 114 QLGQRLGG 121

RESULT 14
US-07-942-245-287
; Sequence 287, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: S30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 287:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-287

Query Match          1.6%; Score 7; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 427 GKSGSPGQ 433
DB 9 GKSGSPGQ 15

RESULT 15
US-07-942-245-325
; Sequence 325, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.

```



;; COUNTRY: United States  
;; ZIP: 20037-3202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: HP 9000/700 Workstation  
;; OPERATING SYSTEM: UNIX  
;; SOFTWARE: In house  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/942,245  
;; FILING DATE: 09-SEP-1992  
;; CLASSIFICATION: 530  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 293-7060  
;; TELEFAX: (202) 293-7860  
;; TELEX: 6491103  
;; INFORMATION FOR SEQ ID NO: 325:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 26 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-07-942-245-325

Query Match 1.6%; Score 7; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 427 GKGSFGQ 433  
| | | | |  
Db 9 GKGSFGQ 15

Search completed: November 24, 2004, 13:57:06  
Job time : 40 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: November 24, 2004, 13:56:29 ; Search time 143 Seconds

(without alignments)  
1093.849 Million cell updates/sec

Title: US-09-513-151a-63

Perfect score: 441

Sequence: 1 MASVAARAVPVGSLRGLQ.....NKEPKGSGQNDLKCSV 441

Scoring table: OLIGO  
Gapex 60.0 , Gapext 60.0

Searched: 1575965 seqs, 354694765 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10F\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	310	70.3	467	14	US-10-097-340-153
2	274	62.1	465	14	Sequence 153, App
3	120	27.2	222	9	US-10-097-340-151
4	120	27.2	222	9	Sequence 151, App
5	48	10.9	57	9	US-09-764-853-586
6	9	2.0	98	16	Sequence 586, App
7	9	2.0	171	16	Sequence 784, App
8	9	2.0	298	17	US-09-864-761-42725
9	9	2.0	302	15	Sequence 42725, A
10	9	2.0	365	14	US-10-437-963-170149
11	9	2.0	450	16	Sequence 170149, A
12	9	2.0	713	14	US-10-437-963-111951
13	8	1.8	43	17	Sequence 111951, A
					US-10-425-115-234061
					Sequence 234061, A
					US-10-425-114-60484
					Sequence 60484, A
					US-10-369-493-8141
					Sequence 8141, App
					US-10-437-963-119875
					Sequence 119875, A
					US-10-437-963-113800
					Sequence 13800, A
					US-10-425-115-234826
					Sequence 234826, A

#### ALIGNMENTS

#### RESULT 1

US-10-097-340-153  
; Sequence 153, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAVATKAR  
; APPLICANT: Steve G. KOVATKAR  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; FILE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; PRIOR FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001-09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10

Sequence 30176, A  
Sequence 202077, A  
Sequence 204283, A  
Sequence 285221, A  
Sequence 205205, A  
Sequence 124, App  
Sequence 284930, A  
Sequence 482, App  
Sequence 217504, A  
Sequence 9951, App  
Sequence 55271, A  
Sequence 205140, A  
Sequence 6, Appli  
Sequence 4, Appli  
Sequence 116049, A  
Sequence 12, Appl  
Sequence 11, Appl  
Sequence 4, Appli  
Sequence 5012, Ap  
Sequence 104297, A  
Sequence 8, Appli  
Sequence 5632, Ap  
Sequence 212921, A  
Sequence 283986, A  
Sequence 284706, A  
Sequence 352543, A  
Sequence 501, App  
Sequence 123015, A  
Sequence 201060, A  
Sequence 361200, A  
Sequence 277334, A

```
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-153

Query Match      70.3%; Score 310; DB 14; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.3e-289;
Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASVAAARAVPVGSLRGLQRTLPVLVILGATGKSTLALQLQRLGGEIVSADSMQVY 60
Db 1 MASVAAARAVPVGSLRGLQRTLPVLVILGATGKSTLALQLQRLGGEIVSADSMQVY 60
QY 61 EGLDIIITNKVSAQQRICRHHMISFVDPVLTNTYVDFRNATATLIEDIFARDKIPITVVG 120
Db 61 EGLDIIITNKVSAQQRICRHHMISFVDPVLTNTYVDFRNATATLIEDIFARDKIPITVVG 120
QY 121 GTNYYIESLLWKVLVNTKPEMGTEKVIDRKLVELEKEDGLVLHKRLSQVDPPEMAAKLHPH 180
Db 121 GTNYYIESLLWKVLVNTKPEMGTEKVIDRKLVELEKEDGLVLHKRLSQVDPPEMAAKLHPH 180
QY 181 DKRVARSLOVFEETGISHSEFLHROHTEEGGPGPLGPKFSPNCIILWLHADQAVLDERL 240
Db 181 DKRVARSLOVFEETGISHSEFLHROHTEEGGPGPLGPKFSPNCIILWLHADQAVLDERL 240
QY 241 DKRVDDMLAAGLLEBLRDFHRRYRQKNVSENSODYQHIGIFQSIGKFEFHYLITEGKCTL 300
Db 241 DKRVDDMLAAGLLEBLRDFHRRYRQKNVSENSODYQHIGIFQSIGKFEFHYLITEGKCTL 300
QY 301 ETSNQLLKKG 310
Db 301 ETSNQLLKKG 310

RESULT 2
US-10-097-340-151
; Sequence 151, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
```

```
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-151

Query Match      62.1%; Score 274; DB 14; Length 465;
Best Local Similarity 100.0%; Pred. No. 6.1e-255;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 STLALQLQRLGGEIVSADSMQVYEGLDIITNKVSAQQRICRHHMISFVDPVLTNTYV 96
Db 35 STLALQLQRLGGEIVSADSMQVYEGLDIITNKVSAQQRICRHHMISFVDPVLTNTYV 94
QY 97 DFRNATATLIEDIFARDKIPITVVGNTNYYIESLLWKVLVNTKPEMGTEKVIDRKLVELEK 156
Db 95 DFRNATATLIEDIFARDKIPITVVGNTNYYIESLLWKVLVNTKPEMGTEKVIDRKLVELEK 154
QY 157 EGLVLHKLSQVDPPEMAAKLHPHDKRVARSLOVFEETGISHSEFLHROHTEGGGPLG 216
Db 155 EGLVLHKLSQVDPPEMAAKLHPHDKRVARSLOVFEETGISHSEFLHROHTEGGGPLG 214
QY 217 GPLKFSNPCIILWLHADQAVLDERLDRKRVDDMLAAGLLEBLRDFHRRYRQKNVSENSODYQ 276
Db 215 GPLKFSNPCIILWLHADQAVLDERLDRKRVDDMLAAGLLEBLRDFHRRYRQKNVSENSODYQ 274
QY 277 HGIFQSIGKFEFHYLITEGKCTLTSNQLLKKG 310
Db 275 HGIFQSIGKFEFHYLITEGKCTLTSNQLLKKG 308

RESULT 3
US-09-764-853-586
; Sequence 586, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 586
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-586

Query Match      27.2%; Score 120; DB 9; Length 222;
Best Local Similarity 100.0%; Pred. No. 7.6e-107;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASVAAARAVPVGSLRGLQRTLPVLVILGATGKSTLALQLQRLGGEIVSADSMQVY 60
```

Db 4 MASVAARAVPVGSLRGLQRTPLVVLGATGKSTLALQLQRLGGEIVSADSMQVY 63

Qy 61 EGLDIITNKVSAQQRICRHHMISFVDPPLVNTVTVDFRNRTALIEDIPARDKIPVVG 120

Db 64 EGLDIITNKVSAQQRICRHHMISFVDPPLVNTVTVDFRNRTALIEDIPARDKIPVVG 123

RESULT 4

US-09-764-853-784

Sequence 784, Application US/09764853

Patent No. US20020090672A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P0206

CURRENT APPLICATION NUMBER: US/09/764,853

CURRENT FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 939

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 784

LENGTH: 222

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (124)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (145)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-853-784

Query Match 27.2%; Score 120; DB 9; Length 222;

Best Local Similarity 100.0%; Pred. No. 7.6e-107; Indels 0; Gaps 0;

Matches 120; Conservative 0; Mismatches 0;

Qy 1 MASVAARAVPVGSLRGLQRTPLVVLGATGKSTLALQLQRLGGEIVSADSMQVY 60

Db 4 MASVAARAVPVGSLRGLQRTPLVVLGATGKSTLALQLQRLGGEIVSADSMQVY 63

Qy 61 EGLDIITNKVSAQQRICRHHMISFVDPPLVNTVTVDFRNRTALIEDIPARDKIPVVG 120

Db 64 EGLDIITNKVSAQQRICRHHMISFVDPPLVNTVTVDFRNRTALIEDIPARDKIPVVG 123

RESULT 5

US-09-864-761-42725

Sequence 42725, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wenheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecmice-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

Prior APPLICATION NUMBER: US 60/180,312

Prior FILING DATE: 2000-02-04

Prior APPLICATION NUMBER: US 60/207,456

Prior FILING DATE: 2000-05-26

Prior APPLICATION NUMBER: US 09/632,366

Prior FILING DATE: 2000-08-03

Prior APPLICATION NUMBER: GB 24263.6

Prior FILING DATE: 2000-10-04

Prior APPLICATION NUMBER: US 60/236,359

Prior FILING DATE: 2000-09-27

Prior APPLICATION NUMBER: PCT/US01/00666

Prior FILING DATE: 2001-01-30

Prior APPLICATION NUMBER: PCT/US01/00667

Qy 59 VYEGLDIITNKVSAQQRICRHHMISFVDPPLVNTVTVDFRNRTALI 106

Db 1 VYEGLDIITNKVSAQQRICRHHMISFVDPPLVNTVTVDFRNRTALI 48

Query Match 10.9%; Score 48; DB 9; Length 57;

Best Local Similarity 100.0%; Pred. No. 5.5e-38; Indels 0; Gaps 0;

Matches 48; Conservative 0; Mismatches 0;

RESULT 6

US-10-437-963-170149

Sequence 170149, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21 (53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 170149

LENGTH: 98

TYPE: PRT

```
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68500C.1.pap
US-10-437-963-170149

Query Match      2.0%; Score 9; DB 16; Length 98;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      211 GGGPLGGPL 219
Db      3 GGGPLGGPL 11

RESULT 7
US-10-437-963-111951
; Sequence 111951, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111951
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(171)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15883C.1.pap
US-10-437-963-111951

Query Match      2.0%; Score 9; DB 16; Length 171;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      211 GGGPLGGPL 219
Db      65 GGGPLGGPL 73

RESULT 8
US-10-425-115-234061
; Sequence 234061, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 234061
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Zea mays

; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4577_145054C.1.pap
US-10-425-115-234061

Query Match      2.0%; Score 9; DB 17; Length 298;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ASVAAARAV 10
Db      242 ASVAAARAV 250

RESULT 9
US-10-425-114-60484
; Sequence 60484, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60484
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3594-059-B4_FLI.pap
US-10-425-114-60484

Query Match      2.0%; Score 9; DB 15; Length 302;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ASVAAARAV 10
Db      246 ASVAAARAV 254

RESULT 10
US-10-369-493-8141
; Sequence 8141, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8141
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-10-369-493-8141

Query Match      2.0%; Score 9; DB 14; Length 365;
Best Local Similarity 100.0%; Pred. No. 11;
```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASVAAARAV 10  
| | | | | | | |  
Db 85 ASVAAARAV 93

## RESULT 11

US-10-437-963-119875  
; Sequence 119875, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 119875  
; LENGTH: 450  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_23049C.1.pep  
US-10-437-963-119875

Query Match 2.0%; Score 9; DB 16; Length 450;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSADSMQVY 60  
| | | | | | | |  
Db 64 VSADSMQVY 72

## RESULT 12

US-10-156-761-13800  
; Sequence 13800, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 13800  
; LENGTH: 713  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13800

Query Match 2.0%; Score 9; DB 14; Length 713;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LGATGTGKS 37  
| | | | | | | |  
Db 47 LGATGTGKS 55

## RESULT 13

US-10-425-115-234826  
; Sequence 234826, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 234826  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_145744C.1.pep  
US-10-425-115-234826

Query Match 1.8%; Score 8; DB 17; Length 43;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 GGPLGGPL 219  
| | | | | | | |  
Db 35 GGPLGGPL 42

## RESULT 14

US-10-029-386-30176  
; Sequence 30176, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE  
; FILE REFERENCE: AEWICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 30176  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR19.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
; OTHER INFORMATION: SWISSPROT HIT: P16258, EVALUE 4.90e-01  
US-10-029-386-30176

Query Match 1.8%; Score 8; DB 14; Length 68;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GSGLRGLQ 20  
 |||||  
 Db 48 GSGLRGLQ 55

## RESULT 15

US-10-425-115-202077  
 ; Sequence 202077, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 202077  
 ; LENGTH: 70  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_115882C.1.pep  
 US-10-425-115-202077

Query Match 1.8%; Score 8; DB 17; Length 70;  
 Best Local Similarity 100.0%; Pred.No. 23;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 GGPLGGPL 219  
 |||||  
 Db 58 GGPLGGPL 65

Search completed: November 24, 2004, 14:10:50  
 Job time : 155 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 24, 2004, 13:52:14 ; Search time 40 seconds  
(without alignments)  
1060.790 Million cell updates/sec

Title: US-09-513-151A-63  
Perfect score: 441  
Sequence: 1 MASVAARAVFVSGSLRGLQ.....NKEFKGSGQNDLKCSV 441

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR.79:\*  
1: Pirl:\*  
2: Pirl2:\*  
3: Pirl3:\*  
4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 S54702	twitching motility
2	9	2.0	428	2 S67176	tRNA isopentenyltr
3	9	2.0	434	2 T38664	tRNA isopentenyltr
4	9	2.0	698	2 B87082	excinnuclease ABC s
5	9	2.0	698	2 G70559	probable uvrB prot
6	8	1.8	125	2 B53380	cell division init
7	8	1.8	162	2 C75437	hypothetical prote
8	8	1.8	216	2 E81670	cytidylate kinase
9	8	1.8	216	2 G71512	probable cmp kinas
10	8	1.8	216	2 T35881	thymidine kinase (
11	8	1.8	246	2 F72464	hypothetical prote
12	8	1.8	263	2 G84083	ABC transporter (A
13	8	1.8	264	1 F64152	hypothetical prote
14	8	1.8	330	2 T48100	tRNA isopentenyl t
15	8	1.8	357	2 F96708	hypothetical prote
16	8	1.8	477	2 C59098	hypothetical prote
17	8	1.8	498	2 B70759	trwB protein - Esc
18	8	1.8	507	2 S43877	NPL4 protein - yea
19	8	1.8	580	2 S34340	excinnuclease ABC c
20	8	1.8	655	2 A64208	excinnuclease ABC c
21	8	1.8	657	2 S73946	excinnuclease ABC c
22	8	1.8	658	2 A11385	excinnuclease ABC c
23	8	1.8	658	2 AC1761	excinnuclease ABC c
24	8	1.8	660	2 C84099	excinnuclease ABC c
25	8	1.8	661	2 G69729	excinnuclease ABC c
26	8	1.8	663	2 G89848	excinnuclease ABC su
27	8	1.8	665	2 A11947	excinnuclease ABC c
28	8	1.8	666	2 C82932	excinnuclease ABC s
29	8	1.8	669	2 S74391	excinnuclease ABC c

RESULT 1

S54702  
twitching motility protein PilU PA0396 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 06-Sep-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S54702; B83595  
R;Whitchurch, C.B.; Mattick, J.S.  
Mol. Microbiol. 13, 1079-1091, 1994  
A;Title: Characterization of a gene, pilU, required for twitching motility but not phase  
A;Reference number: S54702; MUID:95157252; PMID:7854122  
A;Accession: S54702  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-382 <WHI>  
A;Cross-references: UNIPROT:O51532; EMBL:L27667; NID:G443685; PIDN:AAA25965.1; PID:G44368  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Harbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: B83595  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-382 <STO>  
A;Cross-references: GB:AE004477; GB:AE004091; NID:G9946248; PIDN:AAG03785.1; GSPDB:GN0013  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: pilU; PA0396  
C;Superfamily: twitching motility protein pilT

Query Match 2.0%; Score 9; DB 2; Length 382;  
Best Local Similarity 100.0%; Pred.No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 GATGTGKST 38  
|||  
Db 130 GATGTGKST 138

RESULT 2

S67176  
tRNA isopentenyltransferase (EC 2.5.1.8) - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein O5447w; protein YOR274w  
C;Species: Saccharomyces cerevisiae  
C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: S67176; A26717; S72045  
R;Cheret, G.; Sor, F.  
submitted to the Protein Sequence Database, July 1996  
A;Reference number: S67169  
A;Accession: S67176  
A;Molecule type: DNA

A;Residues: 1-428 <CH>  
A;Cross-references: UNIPROT:P07884; EMBL:275182; NID:G1420613; PID:G1420614  
A;Experimental source: strain S288C  
R;Najarian, D.; Dihanich, M.B.; Martin, N.C.; Hopper, A.K.  
Mol. Cell. Biol. 7, 185-191, 1987  
A;Title: DNA sequence and transcript mapping of MOD5: features of the 5' region which su  
A;Reference number: A26717; MUID:8712703; PMID:3031457  
A;Accession: A26717  
A;Molecule type: DNA  
A;Residues: 1-374, 'R', 376-428 <NAJ>  
A;Cross-references: EMBL:M15991  
R;Cheret, G.; Bernardi, A.; Sor, F.  
Yeast 12, 1059-1064, 1996  
A;Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of Saccharomyces  
A;Reference number: S72039; MUID:97051594; PMID:8896271  
A;Accession: S72045  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-428 <CH>  
A;Cross-references: EMBL:X89633; NID:G1279694; PIDN:CAA61780.1; PID:G1419759  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
C;Genetics:  
A;Gene: SGD:MOD5  
A;Cross-references: MIPS:YOR274W; SGD:S0005800  
A;Map position: 15R  
C;Keywords: transferase

Query Match 2.0%; Score 9; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.1; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 KPIPVVGGT 122  
|||||  
DB 104 KPIPVVGGT 112

RESULT 3  
T38664  
trna isopentenyltransferase - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T38664  
R;Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, August 1997  
A;Reference number: Z21804  
A;Accession: T38664  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-434 <MUR>  
A;Cross-references: UNIPROT:Q9UW75; EMBL:AL109739; NID:e1534774; PIDN:CAB52278.1; GSPDB:  
A;Experimental source: strain 972h; cosmid c343  
C;Genetics:  
A;Gene: SPDB:SPAC343.15  
A;Map position: 1

Query Match 2.0%; Score 9; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 1.1; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 KPIPVVGGT 122  
|||||  
DB 93 KPIPVVGGT 101

RESULT 4  
E87082  
excinuclease ABC subunit B [imported] - Mycobacterium leprae  
C;Species: Mycobacterium leprae  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: E87082  
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001  
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
A;Title: Massive gene decay in the leprosy bacillus.  
A;Reference number: A86909; MUID:21128732; PMID:11234002  
A;Accession: E87082  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-698 <STO>  
A;Cross-references: UNIPROT:P57991; GB:AL450380; NID:G13093274; PIDN:CAC31768.1; GSPDB:GH  
C;Genetics:  
A;Gene: uvvB  
C;Superfamily: excinuclease ABC chain B

Query Match 2.0%; Score 9; DB 2; Length 698;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LGATGTGKS 37  
|||||  
DB 40 LGATGTGKS 48

RESULT 5  
G70559  
probable uvvB protein - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C;Accession: G70559  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Suistson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome s  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: G70559  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-698 <COL>  
A;Cross-references: UNIPROT:O06150; GB:Z95554; GB:AL123456; NID:G3261771; PIDN:CAB08886.1  
A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: uvvB  
C;Superfamily: excinuclease ABC chain B  
C;Keywords: ATP; nucleotide binding; P-loop  
F;41-48/Region: nucleotide-binding motif A (P-loop)  
F;335-340/Region: nucleotide-binding motif B  
F;339-342/Region: DEXH motif

Query Match 2.0%; Score 9; DB 2; Length 698;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LGATGTGKS 37  
|||||  
DB 40 LGATGTGKS 48

RESULT 6  
B53380  
cell division initiation protein divIC - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 09-Jul-2004  
C;Accession: B53380; S66092; A69616  
R;Levin, P.A.; Losick, R.  
J. Bacteriol. 176, 1451-1459, 1994  
A;Title: Characterization of a cell division gene from Bacillus subtilis that is require  
A;Reference number: A53380; MUID:94156852; PMID:8113187  
A;Accession: B53380  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-125 <LEV>  
A;Cross-references: UNIPROT:P37471; GB:L23497; NID:G469178; PIDN:AAB38379.1; PID:G385177

R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.  
 DNA Res. 1, 1-14, 1994  
 A;Title: Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome  
 A;Reference number: S65967; MUID:96051385; PMID:7594024  
 A;Accession: S66092  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-125 <OGA>  
 A;Cross-references: EMBL:D26185; NID:g467326; PIDN:BAA05297.1; PID:g467451  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehrlich, S.D.; Emerson, P.T.; Etienne, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallenbach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Mathers, Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroter, R.; Scofield, P.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Winters, M.; Yamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Whittier, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: A69616  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-125 <KUN>  
 A;Cross-references: GB:299104; GB:AL009126; NID:g2632267; PIDN:CAB11838.1; PID:g2632329  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: divIC  
 A;Start codon: TTG  
 C;Superfamily: *Bacillus subtilis* cell division initiation protein divIC

Query Match 1.8%; Score 8; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 3.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 395 NQLKKRRR 402  
 Db 24 NQLKKRRR 31  
 |||||

RESULT 7  
 C75497  
 Hypothetical protein - *Deinococcus radiodurans* (strain R1)  
 C;Species: *Deinococcus radiodurans*  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: C75497  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A;Reference number: A75250; MUID:20036896; PMID:10567266  
 A;Accession: C75497  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-162 <WHI>  
 A;Cross-references: UNIPROT:Q9RWQ7; GB:AE001919; GB:AE000513; NID:g6458307; PIDN:AAF1019  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR0609  
 A;Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 4.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 GTGKSTLA 40  
 |||||

RESULT 8  
 E81670  
 cytidylate kinase TC0737 [imported] - *Chlamydia muridarum* (strain Nigg)  
 C;Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C;Accession: E81670  
 R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, E. C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
 A;Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.  
 A;Reference number: A81500; MUID:20150255; PMID:10684935  
 A;Accession: E81670  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-216 <TET>  
 A;Cross-references: UNIPROT:Q9PUU0; GB:AE002342; GB:AE002160; NID:g7190763; PIDN:AAF3954;  
 A;Experimental source: strain Nigg (MoPn)  
 C;Genetics:  
 A;Gene: TC0737  
 C;Superfamily: cytidylate kinase cmk

Query Match 1.8%; Score 8; DB 2; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 6.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 GTGKSTLA 40  
 |||||  
 Db 10 GTGKSTLA 17

RESULT 9  
 G71512  
 probable cmp kinase - *Chlamydia trachomatis* (serotype D, strain UM3/Cx)  
 C;Species: *Chlamydia trachomatis*  
 C;Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
 C;Accession: G71512  
 R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A;Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis* serotype D.  
 A;Reference number: A71570; MUID:9900809; PMID:9784136  
 A;Accession: G71512  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-216 <ARN>  
 A;Cross-references: UNIPROT:O84458; GB:AE001319; GB:AE001273; NID:g3328881; PIDN:AAC68052  
 A;Experimental source: serotype D, strain UW-3/Cx  
 C;Genetics:  
 A;Gene: cmk  
 C;Superfamily: cytidylate kinase cmk

Query Match 1.8%; Score 8; DB 2; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 6.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 GTGKSTLA 40  
 |||||  
 Db 10 GTGKSTLA 17

RESULT 10  
 T35881  
 thymidine kinase (EC 2.7.1.21) [similarity] - *Streptomyces coelicolor*  
 C;Species: *Streptomyces coelicolor*  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: T35881  
 R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, November 1997  
 A;Reference number: Z21592  
 A;Accession: T35881  
 A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-216 <OLI>  
 A:Cross-references: UNIPROT:O50519; EMBL:AL009204; PIDN:CAA15802.1; GSPDB:GN000070; SCOPED  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOPED:SC9B10.12  
 C:Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 6.2; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 35 GKSTLALQ 42  
 DB 14 GKSTLALQ 21

RESULT 11  
 F72464  
 Hypothetical protein APE2361 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 12-Jul-2004  
 C:Accession: F72464  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatikawa, Y.; Jin-no, K.; Takah  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: F72464  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-246 <KAN>  
 A:Cross-references: UNIPROT:Q9Y9C6; DBJ:AP0000064; NID:G5105945; PIDN:BAA81374.1; PID:G5  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2361  
 C:Superfamily: tropinesterase

Query Match 1.8%; Score 8; DB 2; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 7; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 248 LAAGLLEE 255  
 DB 193 LAAGLLEE 200

RESULT 12  
 G84083  
 ABC transporter (ATP-binding protein) BH3471 [imported] - Bacillus halodurans (strain C-  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 16-Aug-2004  
 C:Accession: G84083  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: G84083  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-263 <STO>  
 A:Cross-references: UNIPROT:Q9K798; GB:AP001518; GB:BA000004; NID:G10175782; PIDN:BA071  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: BH3471  
 C:Superfamily: ATP-binding cassette homology

Query Match 1.8%; Score 8; DB 2; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 7.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 33 GTGKSTLA 40

Db 41 GTGKSTLA 48  
 |||||  
 |||||

RESULT 13  
 F64152  
 Hypothetical protein HI0432 - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: F64152  
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, C  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630; PMID:7543800  
 A:Accession: F64152  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-264 <TIGR>  
 A:Cross-references: UNIPROT:P44710; GB:U32726; GB:L42023; NID:G1573399; PIDN:AAC22091.1;  
 C:Superfamily: hypothetical protein HI0432; mutt domain homology  
 F:158-190/Domain: mutt domain homology <MUT>

Query Match 1.8%; Score 8; DB 1; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 7.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 191 VFEETGIS 198  
 DB 179 VFEETGIS 186

RESULT 14  
 T48100  
 tRNA isopentenyl transferase-like protein - Arabidopsis thaliana  
 N:Alternate names: protein T20010.210  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
 C:Accession: T48100  
 R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lemc  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24484  
 A:Accession: T48100  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-330 <OBE>  
 A:Cross-references: UNIPROT:Q9LYB1; EMBL:AL163816  
 A:Experimental source: cultivar Columbia; BAC clone T20010  
 C:Genetics:  
 A:Map position: 3  
 A:Note: T20010.210

Query Match 1.8%; Score 8; DB 2; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 9.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 30 GATGTGKS 37  
 DB 42 GATGTGKS 49

RESULT 15  
 F96708  
 Hypothetical protein T26J14.3 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: F96708  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Huizlar, L.  
 ansen, N.F.; Hughes, B.; Huizlar, L.  
 Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: F96708  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-357 <STO>  
 A;Cross-references: UNIPROT:Q9CA35; GB:AE005173; NID:96553935; PIDN:AAF16599.1; GSPDB:CN  
 C;Genetics:  
 A;Gene: T26J14.3  
 A;Map position: 1  
 C;Superfamily: delta(2)-isopentenylpyrophosphate transferase

Query Match 1.8%; Score 8; DB 2; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 9.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 VVILGATG 33  
 |||||  
 Db 68 VVILGATG 75

Search completed: November 24, 2004, 13:53:03  
 Job time : 42 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 24, 2004, 13:52:14 ; Search time 192 Seconds  
(without alignments)

1321.564 Million cell updates/sec

Title: US-09-513-151A-63

Perfect score: 441

Sequence: 1 MASVAARAVPVSGSLRGLQ.....NKEPKGKSGQNDLKC SV 441

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	310	70.3	467	1 MOD5_HUMAN	Q93h1 homo sapien
2	98	22.2	324	2 Q6IAC9	Q6IAC9 homo sapien
3	98	22.2	324	2 CAG33507	CAG33507 homo sapi
4	98	22.2	379	2 AAH10741	AAH10741 homo sapi
5	55	12.5	467	1 MOD5_MOUSE	Q80un9 mus musculu
6	55	12.5	467	2 BAB22853	BAB22853 mus muscu
7	46	10.4	208	2 Q8CLM4	Q8CLM4 mus musculu
8	16	3.6	477	2 Q8IMV0	Q8IMV0 drosophila
9	10	2.3	418	2 Q7QEH9	Q7QEH9 anopheles g
10	9	2.0	298	2 Q6WLH1	Q6WLH1 zea mays (m
11	9	2.0	298	2 AAQ62069	AAQ62069 zea mays
12	9	2.0	376	2 Q8P759	Q8P759 xanthomonas
13	9	2.0	376	2 Q8PII4	Q8PII4 xanthomonas
14	9	2.0	381	2 Q8RLD5	Q8RLD5 pseudomonas
15	9	2.0	382	2 Q51532	Q51532 pseudomonas
16	9	2.0	382	2 Q7DCM4	Q7DCM4 pseudomonas
17	9	2.0	383	2 Q6FDQ1	Q6FDQ1 acinetobact
18	9	2.0	428	1 MOD5_YEAST	P0784 saccharomyc
19	9	2.0	431	2 Q8CUJ2	Q8CUJ2 kluyveromyc
20	9	2.0	434	2 Q8UT75	Q8UT75 schizosacch
21	9	2.0	439	2 Q8S926	Q8S926 oryza sativ
22	9	2.0	677	2 Q83MY8	Q83MY8 tropheryma
23	9	2.0	677	2 Q83N17	Q83N17 tropheryma
24	9	2.0	681	2 Q6NH13	Q6NH13 corynebacte
25	9	2.0	681	2 CAE49674	CAE49674 corynebac
26	9	2.0	698	1 UVRB_MYCLE	P57911 mycobacteri
27	9	2.0	698	1 UVRB_MYCTU	Q66150 mycobacteri
28	9	2.0	701	2 Q8FTR2	Q8FTR2 corynebacte
29	9	2.0	702	2 Q7NH00	Q7NH00 gloeobacter
30	9	2.0	712	2 Q8CK11	Q8CK11 streptomyce
31	9	2.0	713	2 Q829Y6	Q829Y6 streptomyce

#### RESULT 1

ID	MOD5_HUMAN	STANDARD;	PRT;	467 AA.
AC	Q9H3H1; Q96FJ3; Q96L45; Q9NXT7;			
DT	05-JUL-2004 (Rel. 44, Created)			
DT	05-JUL-2004 (Rel. 44, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	tRNA isopentenyltransferase, mitochondrial precursor (EC 2.5.1.8)			
DE	(Isopentenyl-diphosphate:RNA isopentenyltransferase) (IIP			
DE	transferase) (IIPase) (IIPPT) (HGROL).			
GN	Name=TRIT1; Synonyms=IPT, MOD5;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=20564178; PubMed=11111046;			
RA	Golovko A., Hjalms G., Sitbon F., Nicander B.;			
RT	"Cloning of a human tRNA isopentenyl transferase.";			
RL	Gene 258:85-93(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RT	Peters J.L., Yan Q., Guan M.X.;			
RL	"Human MOD5 cDNA sequence.";			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Colon;			
RX	PubMed=14702039; DOI=10.1038/ngl285;			
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Oiyashiki M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			
RA	Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,			
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,			
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,			
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,			
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,			
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y.,			
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,			
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,			
RA	Komura Y., Togiya S., Konai P., Hara R., Takeuchi K., Arita M.,			
RA	Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,			
RA	Moriya S., Momiya N., Satoh N., Takami S., Terashima Y., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takenoto M., Kawakami B.,			
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,			
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,			
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,			
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,			

#### ALIGNMENTS

Q740L5 mycobacteri  
Aa03652 mycobacte  
Q9GY56 mus musculu  
Q9F525 escherichia  
P37471 bacillus su  
Q8C2G4 mus musculu  
Q9RWQ7 deinococcus  
Q9RYH0 streptomyce  
Q9F532 escherichia  
O46323 gracilaria  
Q9PJO0 chlamydia m  
O84458 chlamydia t  
O50519 streptomyce  
Q82KI9 streptomyce

32 9 2.0 722 2 Q740L5  
33 9 2.0 722 2 Aa03652  
34 9 2.0 783 2 Q9GY56  
35 9 2.0 785 2 Q9F525  
36 8 1.8 125 1 D1VC\_BACSU  
37 8 1.8 128 2 Q8C2G4  
38 8 1.8 162 2 Q9RWQ7  
39 8 1.8 171 2 Q9RYH0  
40 8 1.8 172 2 Q9F532  
41 8 1.8 194 2 O46323  
42 8 1.8 216 1 KCY\_CHLMU  
43 8 1.8 216 1 KCY\_CHLTR  
44 8 1.8 216 1 KITH\_STRCO  
45 8 1.8 216 2 Q82KI9

RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S., human  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.",  
 RL Nat. Genet. 36:40-45 (2004).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Uterus;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.N.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RP SEQUENCE OF 7-467 FROM N.A. (ISOFORM 4).  
 RX MEDLINE=21444833; PubMed=11560893;  
 RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,  
 RA Barnes T., Hekimi S.;  
 RT "Regulation of physiological rates in *Caenorhabditis elegans* by a  
 RT tRNA-modifying enzyme in the mitochondria.",  
 RL Genetics 159:147-157 (2001).  
 CC -!- FUNCTION: Responsible for the modification of A37 to isopentenyl  
 CC A37 of both cytosolic and mitochondrial tRNAs.  
 CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate +  
 CC tRNA containing 6-isopentenyladenosine.  
 CC -!- PATHWAY: Biosynthesis of the modified base isopentenyladenosine in  
 CC tRNAs.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial (isoforms 1 and 4);  
 CC cytoplasmic and nuclear (isoforms 2 and 3) (potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1;  
 CC IsoId=Q9H3H1-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9H3H1-2; Sequence=VSP\_010719;  
 CC Name=3;  
 CC IsoId=Q9H3H1-3; Sequence=VSP\_010719, VSP\_010720;  
 CC Name=4;  
 CC IsoId=Q9H3H1-4; Sequence=VSP\_010721;  
 CC -!- SIMILARITY: Belongs to the IPP transferase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF074918; AAC31324.1; -;  
 CC EMBL; AY303390; AAP60111.1; -;  
 CC EMBL; AK000068; BAA90923.1; -;  
 CC EMBL; BC010741; AAH10741.1; -;  
 CC EMBL; AY052768; AAL14107.1; -;  
 CC Genew; HGNC:20286; TRIT1.

DR GO: 0004811; F:RNA isopentenyltransferase activity; TAS.  
 DR GO: 0008033; P:RNA processing; NAS.  
 DR InterPro: IPR002627; IPPT.  
 DR InterPro: IPR007087; Znf\_C2H2.  
 DR Pfam: PF01715; IPPT; 1. C2H2.  
 DR ProDom: PD004674; IPPT; 1.  
 DR TIGRFAMs: TIGR00174; miaA; 1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN 1.  
 KW Alternative splicing; ATP-binding; Mitochondrion; Nuclear protein;  
 KW Transferase; Transient peptide; tRNA processing; Zinc-finger.  
 FT TRANSIT 1 47 tRNA isopentenyltransferase.  
 FT CHAIN 48 467  
 FT ZNF\_FING 395 419  
 FT VARSPPLIC 1 141 Missing (in isoform 2 and isoform 3).  
 FT VARSPPLIC 235 236 Missing (in isoform 3).  
 FT VARSPPLIC 311 336 Missing (in isoform 4).  
 FT CONFLICT 446 446 H -> Y (in Ref. 5).  
 FT SEQUENCE 467 AA; 5275 MW; 634469919D7F56A5 CRC64;  
 SQ  
 Query Match 70.3%; Score 310; DB 1; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-311;  
 Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MASVAAARAVPGVSGRLGRLQRLPLVWILGATGCTKSTLALQLGRLGGEIVSADSMOYV 60  
 DB 1 MASVAAARAVPGVSGRLGRLQRLPLVWILGATGCTKSTLALQLGRLGGEIVSADSMOYV 60  
 QY 61 EGLDITNKVSAEQRICEHHMISFVDPVLTNTYTVDFRNATALIEDIFARDKIPVVG 120  
 DB 61 EGLDITNKVSAEQRICEHHMISFVDPVLTNTYTVDFRNATALIEDIFARDKIPVVG 120  
 QY 121 GNNYIESLLKVLVNTKQPMGTEKVDKRVLEKEDGLVHLKLSQVDPMAAKLHPH 180  
 DB 121 GNNYIESLLKVLVNTKQPMGTEKVDKRVLEKEDGLVHLKLSQVDPMAAKLHPH 180  
 QY 181 DKRKVARSLQVEETGISHSEFLHQHTEEGGGPGGPGPKFSNPCLILWHLADQVLDRL 240  
 DB 181 DKRKVARSLQVEETGISHSEFLHQHTEEGGGPGGPGPKFSNPCLILWHLADQVLDRL 240  
 QY 241 DKRVDDMLAAGLLELRDPHRRYNNQKVSNSQDYQHGFQSIGKPEFHEYLITEGKCTL 300  
 DB 241 DKRVDDMLAAGLLELRDPHRRYNNQKVSNSQDYQHGFQSIGKPEFHEYLITEGKCTL 300  
 QY 301 ETSNQLKKG 310  
 DB 301 ETSNQLKKG 310  
 RESULT 2  
 Q6IAC9 ID Q6IAC9 PRELIMINARY; PRT; 324 AA.  
 AC Q6IAC9;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE TRIT1 protein.  
 GN Name=TRIT1;  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the IPP transferase family.  
 CC EMBL; CR457226; CAG33507.1; -;  
 DR InterPro: IPR002627; IPPT.  
 DR InterPro: IPR007087; Znf\_C2H2.  
 DR InterPro: IPR003604; Znf\_U1.



```

DR Pfam; PF01715; IPPT; 1.
DR ProDom; PD004674; IPPT; 1.
DR SMART; SM00451; Znf U1; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW ATP-binding; Transferase.
SQ SEQUENCE 324 AA; 37193 MW; AB6C258689D9BE70 CRC64;

Query Match 22.2%; Score 98; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 3.4e-92;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 PGPIVPPVYGLVSDVSKWESVLEPALEIVQSFIOGHKPTATPIKMPYNEAENKRSYHL 370
Db 194 PGPIVPPVYGLVSDVSKWESVLEPALEIVQSFIOGHKPTATPIKMPYNEAENKRSYHL 253
Qy 371 CDLCDRIIIGDREWAHAKSKSHLNQLKRRRLSDAV 408
Db 254 CDLCDRIIIGDREWAHAKSKSHLNQLKRRRLSDAV 291

RESULT 3
CAG33507 PRELIMINARY; PRT; 324 AA.
AC CAG33507;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE TRIT1 protein.
GN TRIT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schattner R., Henze S., Korn B.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
vector (pDONR201)".;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR457226; CAG33507.1; -.
SQ SEQUENCE 324 AA; 37193 MW; AB6C258689D9BE70 CRC64;

Query Match 22.2%; Score 98; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 3.4e-92;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 PGPIVPPVYGLVSDVSKWESVLEPALEIVQSFIOGHKPTATPIKMPYNEAENKRSYHL 370
Db 194 PGPIVPPVYGLVSDVSKWESVLEPALEIVQSFIOGHKPTATPIKMPYNEAENKRSYHL 253
Qy 371 CDLCDRIIIGDREWAHAKSKSHLNQLKRRRLSDAV 408
Db 254 CDLCDRIIIGDREWAHAKSKSHLNQLKRRRLSDAV 291

RESULT 4
AAH10741 PRELIMINARY; PRT; 379 AA.
AC AAH10741;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE TRIT1 protein (Fragment).
GN TRIT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA TISSUE=Uterus;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

```

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010741; AAH10741.2; -.
FT NON-TER 1
SQ SEQUENCE 379 AA; 42774 MW; 076ASD6555C537D5 CRC64;

Query Match 22.2%; Score 98; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 4e-92;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 PGPIVPPVYGLVSDVSKWESVLEPALEIVQSFIOGHKPTATPIKMPYNEAENKRSYHL 370
Db 249 PGPIVPPVYGLVSDVSKWESVLEPALEIVQSFIOGHKPTATPIKMPYNEAENKRSYHL 308
Qy 371 CDLCDRIIIGDREWAHAKSKSHLNQLKRRRLSDAV 408
Db 309 CDLCDRIIIGDREWAHAKSKSHLNQLKRRRLSDAV 346

RESULT 5
MOD5 MOUSE STANDARD; PRT; 467 AA.
ID MOD5 MOUSE
AC Q80UN9; Q9D1H5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE tRNA isopentenyltransferase, mitochondrial precursor (EC 2.5.1.8)
DE (isopentenyl-diphosphate:tRNA isopentenyltransferase) (IPPT).
DE transferase (IPTase) (IPPT).
GN Name=Trit1; Synonyms=Ipt;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaio I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

```

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilning L.O., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC STRAIN=Czech II; TISSUE=Breast tumor; and Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kertanen M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinnwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Responsible for the modification of A37 to isopentenyl  
CC A37 of both cytosolic and mitochondrial tRNAs (By similarity).  
CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate +  
CC tRNA containing 6-isopentenyladenosine.  
CC -!- PATHWAY: Biosynthesis of the modified base isopentenyladenosine in  
CC tRNAs.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial (isoform 1), cytoplasmic and  
CC nuclear (isoform 2) (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q80UN9-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q80UN9-2; Sequence=VSP\_010722;  
CC -!- SIMILARITY: Belongs to the IPP transferase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AK003556; BAB22853.2; -  
CC EMBL; BC019812; AAH19812.1; -  
CC EMBL; BC051040; AAH51040.1; -  
CC MGD; MGI:1914216; Tt1t1.  
CC InterPro; IPR002627; IPRP.  
CC InterPro; IPR007087; Znf\_C2H2.  
CC Pfam; PF01715; IPRP; 1.  
CC ProDom; PD004674; IPRP; 1.  
CC TIGRFAMs; TIGR00174; miaA; 1.

DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN 1.  
KW Alternative splicing; ATP-binding; Mitochondrion; Nuclear protein;  
KW Transferase; Transist peptide; tRNA processing; Zinc-finger.  
FT TRANSIT 1 47 Mitochondrion (Potential).  
FT CHAIN 48 467 tRNA isopentenyltransferase.  
FT ZN\_FING 395 419 C2H2-type (Potential).  
FT VARSPLIC 1 141 Missing (in isoform 2).  
FT FTID=VSP\_010722.  
FT CONFLICT 379 379 M -> T (in Ref. 2; AAH51040).  
FT CONFLICT 384 384 A -> T (in Ref. 2; AAH51040).  
FT SEQUENCE 467 AA; 52436 MW; BC0EEF50AB05FD82 CRC64;  
SQ  
Query Match 12.5%; Score 55; DB 1; Length 467;  
Best Local Similarity 100.0%; Pred. No. 1.7e-4; Gaps 0;  
Matches 55; Conservative 0; Mismatches 0; Indels 0;  
OY 21 RTPLPVLVILGATGKSTLALQGLGGEIVSADSMQVYGLDIITNKVSAEQ 75  
DB 21 RTPLPVLVILGATGKSTLALQGLGGEIVSADSMQVYGLDIITNKVSAEQ 75  
RESULT 6  
BAB22853  
ID BAB22853 PRELIMINARY; PRT; 467 AA.  
AC BAB22853;  
DT 14-APR-2004 (TrEMBLrel. 27, Created)  
DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)  
DE 18-day embryo whole body cDNA, RIKEN full-length enriched library,  
DE clone:1110007017 product:tRNA ISOPENTENYLPHOSPHATE TRANSFERASE  
DE homolog.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085860; PubMed=11217851;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20499374; PubMed=11042159;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11078861;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,

```

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK003556; BAB22853.2; -.
KW Transferase.
SQ SEQUENCE 467 AA; 52436 MW; BC0BEF50AE05FD82 CRC64;

Query Match 12.5%; Score 55; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.7e-47;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 RTPLVILGATGKSTLALQGLGGEIVSADSNQVYEGLDIITNKVSAEQ 75
Db 21 RTPLVILGATGKSTLALQGLGGEIVSADSNQVYEGLDIITNKVSAEQ 75

RESULT 7
Q8C1M4 PRELIMINARY; PRT; 208 AA.
AC Q8C1M4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DE library, clone:2310075G14 product:TRNA ISOPENTENYLPROPHOSPHATE
DE TRANSFERASE homolog.
GN Name=Trt1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBT_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).

```

```

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20530913; PubMed=11078661;
RA Shibata K., Itoh M., Aizawa K., Kitsuai T., Tashiro H., Itoh M.,
RA Konno H., Akiyama J., Nishi K., Matsumoto S., Hazama M., Nishine T., Harada A.,
RA Sumi N., Ishii Y., Nakamura S., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the IPP transferase family.
DR EMBL; AK010176; BAC25285.1; -.
DR MGD; MGI:1914216; Trt1.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004811; F:RNA isopentenyltransferase activity; IEA.
DR GO; GO:0008033; P:RNA processing; IEA.
DR InterPro; IPR002627; IPPT.
DR Pfam; PF01715; IPPT; 1.
DR ProDom; PD004674; IPPT; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 208 AA; 23979 MW; 5553F869C83C5098 CRC64;

Query Match 10.4%; Score 46; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 LHKRLSQVDPMAAKLPHDKKVARSLQVFETGISHSEFLRHQ 207
Db 21 LHKRLSQVDPMAAKLPHDKKVARSLQVFETGISHSEFLRHQ 66

RESULT 8
Q8IMV0 PRELIMINARY; PRT; 477 AA.
AC Q8IMV0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG31381-PA.
GN ORFNames=CG31381;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```



```

DR ProDom; PD005388; IPTrans like; 1.
DR TIGR00174; r1aA_1
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW ATP-binding; Transferase.
FT NON TER 418 418
FT NON TER 1
SQ SEQUENCE 418 AA; 47820 MW; B643066A00FAE3BD CRC64;

Query Match 2.3%; Score 10; DB 2; Length 418;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 119 VGGTNYIYES 128
Db 95 VGGTNYIYES 104
|||||

RESULT 10
Q6WLH1 PRELIMINARY; PRT; 298 AA.
AC Q6WLH1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Single myb histone 6.
GN Name=Smh6;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
Tissue=Central region of pre-emerged;
RX MEDLINE=22975137; PubMed=14576282;
RA Marian C.O., Bordoli S.J., Goltz M., Santarella R.A., Jackson L.P.,
RA Danilevskaya O., Beckstette M., Meeley R., Bass H.W.;
RT "The maize single myb histone 1 gene, Smh1, belongs to a novel gene
RT family and encodes a protein that binds telomere DNA repeats in
RT vitro.";
RL Plant Physiol. 133:1336-1350(2003).
DR EMBL; AY280632; AAQ62069.1; -.
SQ SEQUENCE 298 AA; 33004 MW; 706E6E8358AFCB7D CRC64;

Query Match 2.0%; Score 9; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ASVAARAV 10
Db 242 ASVAARAV 250
|||||

RESULT 11
AAQ62069 PRELIMINARY; PRT; 298 AA.
AC AAQ62069;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)

```

```

DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Single myb histone 6.
GN SMH6.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SEQUENCE FROM N.A.
Tissue=Central region of pre-emerged;
RX MEDLINE=22975137; PubMed=14576282;
RA Marian C.O., Bordoli S.J., Goltz M., Santarella R.A., Jackson L.P.,
RA Danilevskaya O., Beckstette M., Meeley R., Bass H.W.;
RT "The maize single myb histone 1 gene, Smh1, belongs to a novel gene
RT family and encodes a protein that binds telomere DNA repeats in
RT vitro.";
RL Plant Physiol. 133:1336-1350(2003).
DR EMBL; AY280632; AAQ62069.1; -.
SQ SEQUENCE 298 AA; 33004 MW; 706E6E8358AFCB7D CRC64;

Query Match 2.0%; Score 9; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ASVAARAV 10
Db 242 ASVAARAV 250
|||||

RESULT 12
Q8P759 PRELIMINARY; PRT; 376 AA.
AC Q8P759;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Twitching motility protein.
GN Name=PilU; OrderedLocusNames=XCC2754;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
SEQUENCE FROM N.A.
RX STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorfi H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferrio M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Spindola L.A.F., Takita M.A., Tamura R.C., de Souza R.F.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AB012388; AAM42026.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0001166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPass.
DR InterPro; IPR001482; GSPT_E.

```

```

DR InterPro; IPR006321; Pili.
DR Pfam; PF00437; GSP11_E; 1.
DR ProDom; PD000739; GSP11_E; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR01420; pili_fam; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 376 AA; 41996 MW; 9A85813AA18863C5 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GATGTGKST 38
DB 133 GATGTGKST 141

RESULT 13
Q8PIH4 PRELIMINARY; PRT; 376 AA.
AC Q8PIH4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Twitching motility protein.
GN Name=pili; OrderedLocNames=XAC2923;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Cannavaro F., Cardoso J., Chambergo P., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.B., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsunaga A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AB011934; AAM37768.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001482; GSP11_E.
DR Pfam; PF00437; GSP11_E; 1.
DR ProDom; PD000739; GSP11_E; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR01420; pili_fam; 1.
KW ATP-binding.
SQ SEQUENCE 381 AA; 42417 MW; C69B6B6C24EA3DB CRC64;

Query Match 2.0%; Score 9; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GATGTGKST 38
DB 133 GATGTGKST 138

RESULT 15
Q51532 PRELIMINARY; PRT; 382 AA.
AC Q51532;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pili protein.
GN Name=pili;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01;
RX MEDLINE=91285432; PubMed=1676395;
RA Whitchurch C.B., Hobbs M., Livingston S.P., Krishnappillai V.,
RA Mattick J.S.;
RT "Characterisation of a Pseudomonas aeruginosa twitching motility gene
RT and evidence for a specialised protein export system widespread in
RT eubacteria";
RL Gene 101:33-44(1991).
RN [2]

```

RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RA Whitechurch C.B., Mattick J.S.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
DR EMEL; L27667; AAA25965.1; -.  
DR PIR; S54702; S54702.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR001482; GSP11\_E.  
DR InterPro; IPR006321; Pilt.  
DR Pfam; PF00437; GSP11\_E; 1.  
DR ProDom; PD000739; GSP11\_E; 1.  
DR SMART; SM00382; AAA; 1.  
DR TIGRFAMs; TIGR01420; pilt\_fam; 1.  
KW ATP-binding.  
SQ SEQUENCE 382 AA; 42532 MW; 0266E5D0FED35E17 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;  
Best Local Similarity 100.0%; Pred.No.6.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 GATGTGKST 38  
| | | | |  
Db 130 GATGTGKST 138

Search completed: November 24, 2004, 13:56:22  
Job time : 195 secs

***This Page Blank (uspto)***